Technical white paper

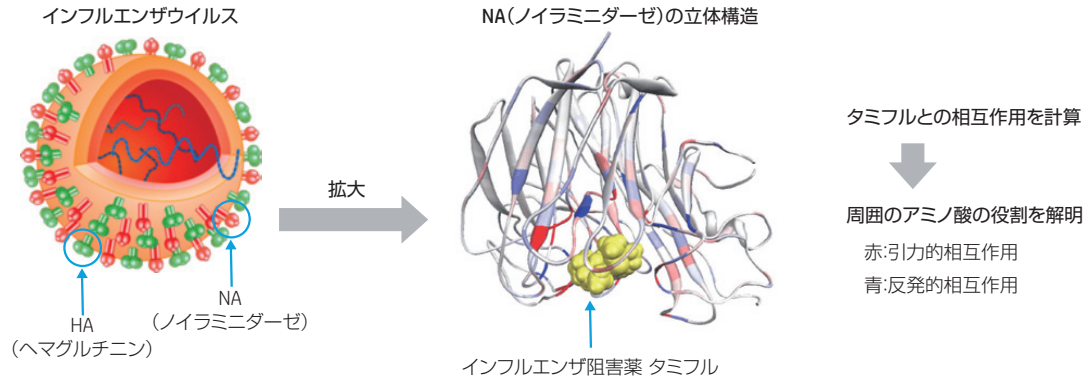
HP ProLiant DL980 G7での MIZUHO/BioStation 2.0の評価

創薬における医薬品化合物の設計においては、医薬品の標的となるタンパク質に、より選択的に、より強く結合する化合物を設計することが重要です。従来の化合物設計には、主に古典力場や経験的なスコア関数を用いたドッキングシミュレーションが用いられてきましたが、近年、量子化学計算に基づいたドラックデザインへの期待が高まっています。MIZUHO/BioStationは、従来、計算することが困難であったタンパク質などの大規模分子の電子状態計算や相互作用解析を可能にしたバイオ分子相互作用シミュレータです。MIZUHO/BioStationは、文部科学省「イノベーションソフトウェアの研究開発プロジェクト（RISS）」において開発されたBioStationを基に、みずほ情報総研（株）が製品化し、独自の機能とサポートを加えて販売しているソフトウェアです。この技術白書では、MIZUHO/BioStationの並列化効率をインテル® Xeon® プロセッサー E7ファミリー8基（80コア）を搭載したHP ProLiant DL980 G7で評価した結果をご紹介します。

1. MIZUHO/BioStationについて

MIZUHO/BioStationは、フラグメント分子軌道（FMO）法に基づく量子化学計算を行うバイオ分子相互作用シミュレータです。FMO法では、分子や分子の集合体を適当なサイズのフラグメント（断片）に分割し、フラグメントモノマーとフラグメントダイマーなどについて分子軌道（MO）計算を行い、分子全体のエネルギーや電子状態を計算します。低分子化合物で成功を収めた分子軌道計算を、精度を落とさずにタンパク質のような巨大分子へ適用することが出来ます。MIZUHO/BioStationには専用GUIであるMIZUHO/BioStation Viewerが付属しており、フラグメント分割や入力パラメータ設定などのプリ処理から、フラグメント間相互作用エネルギー（IFIE）解析、軌道相互作用解析（CAFI, FILM）などのポスト処理までをサポートします。

図1. インフルエンザウイルスタンパク質とタミフルとの相互作用を解析した例



インテル® Xeon® プロセッサー E7 ファミリー

MIZUHO/BioStationは、これまでに、インフルエンザウイルスの膜表面タンパク質における抗原抗体反応や抗ウイルス薬タミフルの酵素ー阻害剤結合、生活習慣病や乳がんにかかわる核内受容体のリガンド結合や転写の分子メカニズム、肺がんの分子標的薬とのドッキング、DNA転写等、さまざまな分子認識における相互作用の解析を行った実績を有しています。図1はインフルエンザウイルスタンパク質とタミフルとの相互作用を解析した例ですが、タミフルと周囲のアミノ酸の相互作用を計算・可視化することでタミフルの結合に重要なアミノ酸を知ることができます。また、タンパク質、DNAなどの生体高分子への適用はもちろん、今後、ナノ材料分子への適用が期待されています。

2. MIZUHO/BioStationの機能概要

① 専用GUIによるフラグメント分割機能および入力パラメータの設定

タンパク質、DNAの自動フラグメント分割はもちろん、マニュアル操作による医薬品候補化合物などの手動分割ができます。また、GUI上で作成した構造について、入力パラメータの設定を行うことができます。

② フラグメント間相互作用エネルギー(IFIE)解析

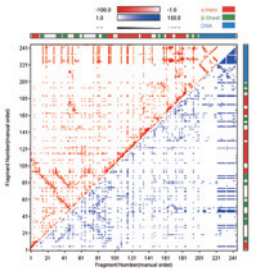
FMO計算の実行により、IFIEと呼ばれる各フラグメント間の相互作用エネルギー値を得ることができます。IFIEを用いて、タンパク質と化合物、タンパク質とタンパク質、DNAとタンパク質、等の分子間相互作用を可視化・解析することが可能となります。

③ 軌道相互作用解析(CAFI、FILM)、グリッドデータ解析

電荷移動相互作用の解析機能CAFIや、CH/πやπ/πなど分散相互作用の解析機能FILMを有しています。これらの機能により、エネルギー解析よりも詳細な軌道相互作用を評価することができます。また、電子密度分布や静電ポテンシャル、フラグメントの分子軌道(MO)などを可視化できます。

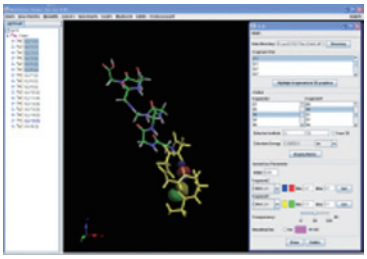
図2. MIZUHO/BioStationの機能概要

IFIEの立体構造表示



IFIE mapの表示

フラグメント間相互作用エネルギー解析



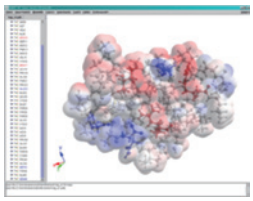
FILMによる軌道間相互作用の解析

軌道間相互作用解析

フラグメント分割処理例

入力パラメータ設定画面

フラグメント分割および入力ファイルの作成



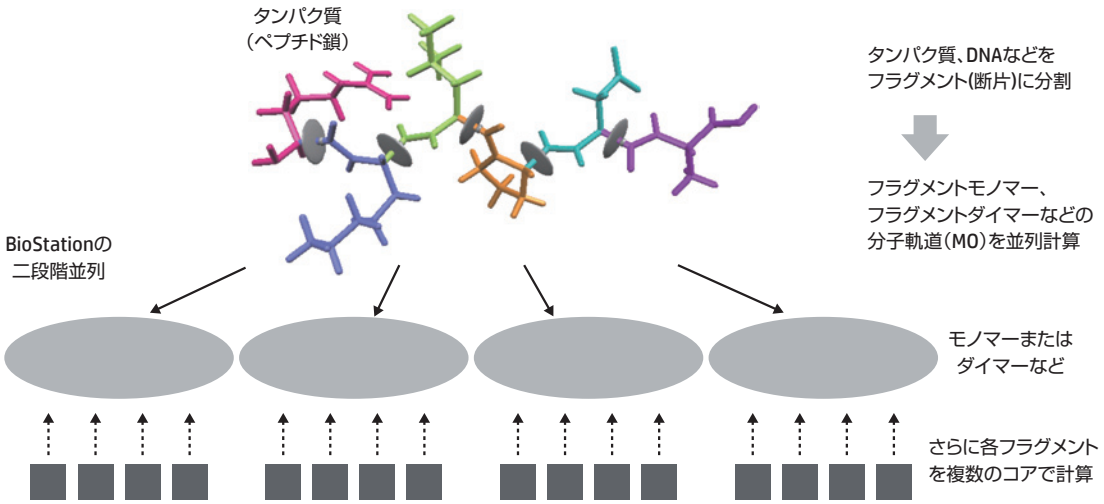
静電ポテンシャルマップの表示

グリッドデータ解析

3. フラグメント分子軌道(FMO)法と並列計算

前述の通りFMO法では、分子や分子の集合体を適当なサイズのフラグメント(断片)に分割し、フラグメントモノマーとフラグメントダイマーなどについて分子軌道(MO)計算を行い、分子全体のエネルギーや電子状態を計算します。複数のフラグメントのMO計算を並列に行うことで、タンパク質のような巨大分子の電子状態を高速に計算することができます。さらにMIZUHO/BioStationでは、各フラグメント内の電子状態計算も並列化されます。こうして複数フラグメントの並列計算と、各フラグメント内の並列計算の二段階の並列化が実装されているため、より高い並列化効率を期待することができます。

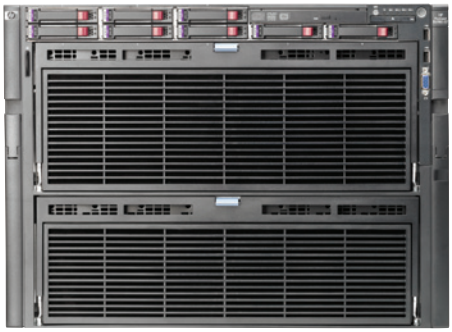
図3. タンパク質(ペプチド鎖)のフラグメント分割と並列化計算の模式図



4. HP ProLiant DL980 G7、Open MPIで環境を構築

今回のベンチマークテストにはインテル® Xeon® プロセッサー E7ファミリー8基(80コア)を搭載したHP ProLiant DL980 G7を使用しました。ソフトウェア環境としては、OSにCentOS、並列化のためのライブラリにはOpen MPIを使用しました。
高次電子相関を考慮した量子化学計算では大きなメモリーを必要とするため、2TBのメモリーが搭載されたHP ProLiant DL980 G7は最適なハードウェア環境でした。また、1つのシステムで80個のコアを利用できるため、複数のシステムにまたがって並列計算を実行する場合に比べて簡単な環境設定で並列計算を実行できました。

図4. ベンチマークのためのシステム環境



ハードウェア:**HP ProLiant DL980 G7**
搭載プロセッサー:インテル® Xeon® プロセッサー E7-2870 8基/80コア
搭載メモリー:2TB
搭載記憶領域:900GB SAS HDD×8台　合計7.2TB
OS:CentOS
ライブラリ:Open MPI

インテル® Xeon® プロセッサー
E7 ファミリー

5. 80コアを使えば、処理時間はシングルコアの約1/40に

HP ProLiant DL980 G7の使用コア数を増やしながら、小規模タンパク質TrpCageのFMO計算を行いました。図5に使用コア数ごとの計算時間（m）を示します。シングルコアでは279.0分（約4時間半）掛かっていた計算時間が、80コアで並列化することによりわずか7.4分と大幅に短縮されています。図6に示した並列化による処理効率でみると80コアを使えば、処理時間はシングルコアの約1/40になることがわかります。

図5. 小規模タンパク質TrpCageのFMO計算(FMO-MP2/6-31G)の計算時間

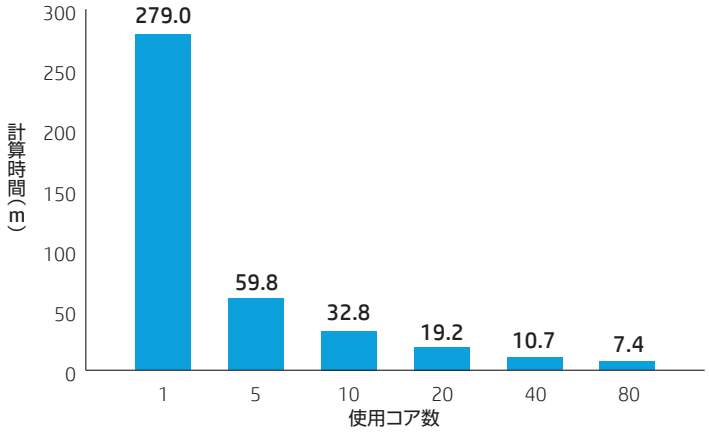
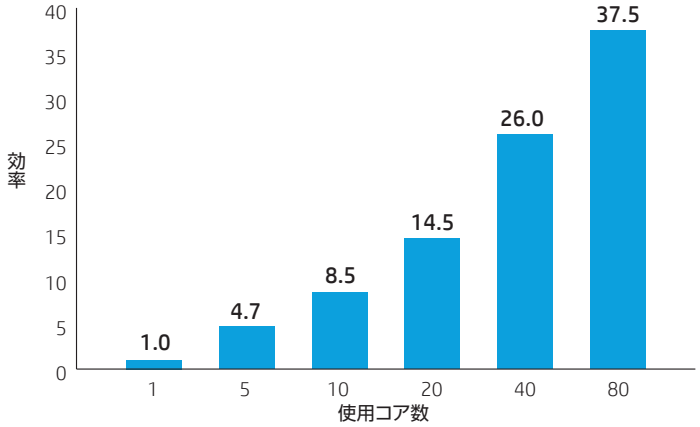


図6. 並列化による処理効率の違い



MIZUHO/BioStationに関して

BioStationは国立大学法人東京大学の登録商標であり、MIZUHO/BioStationはみずほ情報総研の製品となります。

安全に関するご注意　ご使用の際は、商品に添付の取扱説明書をよくお読みの上、正しくお使いください。水、湿気、油煙等の多い場所に設置しないでください。火災、故障、感電などの原因となることがあります。

お問い合わせはカスタマー・インフォメーションセンターへ
03-5749-8328 月～金 9:00～19:00　土 10:00～17:00（日、祝祭日、年末年始および5/1を除く）
機器のお見積もりについては、代理店、または弊社営業にご相談ください。
HP ProLiantに関する情報は http://www.hp.com/jp/proliant

Intel、インテル、Intel ロゴ、Intel Inside、Intel Inside ロゴ、Xeon、Xeon Insideは、アメリカ合衆国および /
またはその他の国におけるIntel Corporationの商標です。
記載されている会社名および商品名は、各社の商標または登録商標です。
記載事項は2013年9月現在のものです。
本カタログに記載されている情報は取材時におけるものであり、閲覧される時点で変更されている可能性があります。あらかじめご了承ください。
© Copyright 2013 Hewlett-Packard Development Company,L.P.

本カタログは、環境に配慮した用紙と植物性大豆油インキを使用しています。 PRINTED WITH SOY INK

日本ヒューレット・パッカード株式会社
〒136-8711 東京都江東区大島2-2-1